

# Towards a Data Interchange Format for XAS and Related Spectroscopies

---

**Bruce Ravel**

Synchrotron Methods Group, Ceramics Division  
Materials Measurement Laboratory  
National Institute of Standards and Technology  
&

Local Contact, Beamline X23A2  
National Synchrotron Light Source

NSLS Users' Meeting  
Data Acquisition for Fluorescence Imaging and Spectroscopy

25 May 2011

---

# Acknowledgements

The people participating on the XASFORMAT mailing list:



Matt Newville (CARS & APS)



Armando Solé (ESRF) (whose picture cannot be found on the web!)



James Hester (ANSTO)



Gerd Wellenreuther (DESY)



Darren Dale (CHESS)

# A single-spectrum XAS data standard

or, is this talk really going to be about such a boring topic?

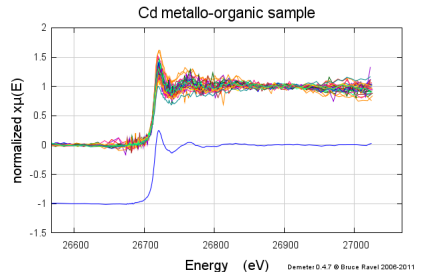
Why, yes! Yes, it is!

A single XAS spectrum is a very useful *unit of currency* for a number of reasons:

- We send data to and receive data from our collaborators
- We publish individual spectra in journal articles
- We extract XAS spectra from large, complex, multi-spectral data sets
- We write web-based and desktop applications (for instance, a standards database) that traffic in single spectra
- We write data analysis software that needs to reliably import data that comes from many sources.

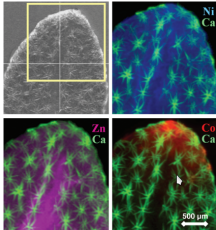
# Conventional XAS

- 1 In a conventional XAS experiment, we measure a sample somewhere between 2 and 10,000 scans, possibly requiring dead-time or other corrections
- 2 Some data processing happens to correct, calibrate, and/or align the data
- 3 Those scans are then merged into a single spectrum that becomes our **unit of currency**

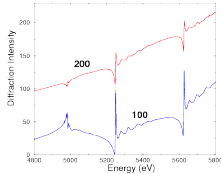


A data interchange standard is about how we express the merged spectrum (i.e. the **blue one**).

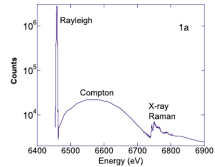
# We do a lot of fancy things these days



We do imaging experiments that sample the heterogeneity of our samples.



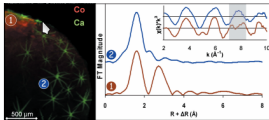
We do anomalous scattering experiments that yield energy-dependent scattering intensities.



We do non-resonant inelastic scattering experiments.

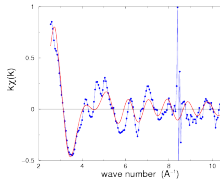
Tappero et al. *New Phytologist* **175**:4, 641-654, (2007) doi:10.1111/j.1469-8137.2007.02134.x  
Ravel et al. *PRB* **60**, 778-785 (1999) doi:10.1103/PhysRevB.60.778  
Bergmann, et al. *Chem. Phys. Lett.* **369** 184 (2003) doi:10.1016/S0009-2614(02)02003-1

# We do a lot of fancy things these days



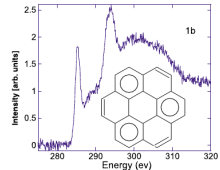
We do imaging experiments that sample the heterogeneity of our samples.

...in this case, a spectrum from a particular spot on the sample.



We do anomalous scattering experiments that yield energy-dependent scattering intensities.

...in this case, the XAS spectrum from a particular crystallographic site.

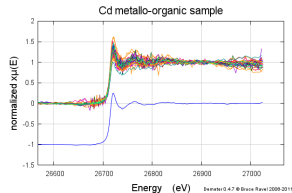


We do non-resonant inelastic scattering experiments.

...in this case, a low energy K edge measured by energy loss.

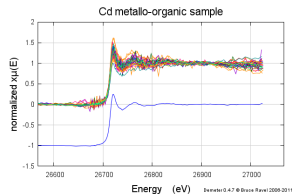
# Beamline data formats

- Every beamline has it's own way of recording data
- Most use ascii files, some use more complex data formats
- Each beamline has good reasons for doing things their own way



# Beamline data formats

- Every beamline has it's own way of recording data
- Most use ascii files, some use more complex data formats
- Each beamline has good reasons for doing things their own way



## NSLS XDAC

```

XDAC V1.4 Datafile V1
"au.b04" created on 3/15/09 at 1:28:27 PM on X-23A2
Diffraction element= Si (311). Ring energy= 2.80 GeV
EO= 11919.00
NUM_REGIONS= 4
SRB= -200 -20 30 60 20k
SRSS= 10 0.25 0.05k 0.05k
SPP= 1 1 1 0.25k
Settling time= 0.30
  
```

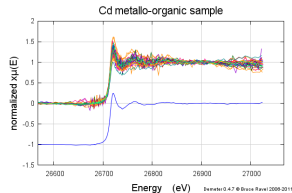
```

Offsets= 122.00 85.78 0.00
Gains= 8.00 8.00 1.00
Au foil, NSLS X23A2, 20% Ar in Io and It
with harmonic rejection mirror
  
```

Energy	IO	It	IntTime
11719.00294	18352.0000	15872.2222	1.0000
11728.99732	18380.0000	15934.2222	1.0000
11739.00126	18381.0000	15980.2222	1.0000
...			

# Beamline data formats

- Every beamline has it's own way of recording data
- Most use ascii files, some use more complex data formats
- Each beamline has good reasons for doing things their own way



## NSLS XDAC

```
XDAC V1.4 Datafile V1
"au.b04" created on 3/15/09 at 1:28:27 PM on X-23A2
Diffraction element= Si (311). Ring energy= 2.80 GeV
EO= 11919.00
NUM_REGIONS= 4
SRB= -200 -20 30 60 20k
SRSS= 10 0.25 0.05k 0.05k
SPP= 1 1 1 0.25k
Settling time= 0.30
```

```
Offsets= 122.00 85.78 0.00
Gains= 8.00 8.00 1.00
Au foil, NSLS X23A2, 20% Ar in Io and It
with harmonic rejection mirror
```

Energy	IO	It	IntTime
11719.00294	18352.0000	15872.2222	1.0000
11728.99732	18380.0000	15934.2222	1.0000
11739.00126	18381.0000	15980.2222	1.0000
...			

## Photon Factory and SPring-8

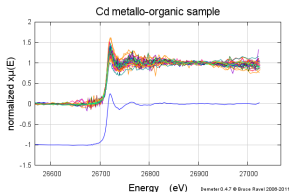
```
9809      KKK-PF      BL12C
G:hgcys-11.001 07.05.12 23:28 - 07.05.12 23:55
Hg:H2Cys 1:2 pH = 12.86, 100 mM, prep. at PF, 5 mm Teflon, stirred 4
Ring : 2.5 GeV 348.8 mA - 342.8 mA
Mono : SI(111) D= 3.13551 A Initial angle= 9.25969 deg
BL12C Transmission( 2) Repetition= 6 Points= 818
Param file : A:hgk16 energy axis(2) Block = 6
```

Block	Init-Eng	final-Eng	Step/eV	Time/s	Num
1	12049.00	12150.00	6.00	1.00	17
2	12150.00	12320.00	.35	1.00	486
3	12320.00	12400.00	1.00	2.00	80
4	12400.00	12600.00	2.50	3.00	80
5	12600.00	13040.00	4.00	3.00	110
6	13040.00	13260.00	5.00	4.00	45

```
Ortec(-1) NDCH = 3
Angle(c) Angle(o) time/s 2 3
Mode 0 0 1 2
Offset 0 0 826.150 652.975
9.44433 9.44420 1.00 252916 592687
9.43958 9.43960 1.00 256349 604260
9.43483 9.43480 1.00 256429 607846
```

# Problems with beamline formats

- They are all different!
- They require additional processing in order to display  $\mu(E)$ , including
  - Conversion to energy
  - Dead-time or other corrections
- Ambiguous metadata, for instance
  - How is the beamline identified?
  - What constitutes a user comment?
  - What describes the condition of the source or the beamline?
- XAS data analysis software and other plotting software may have difficulty importing and interpreting the data
- This data is probably not appropriate for submission to a journal as supplemental material



## Data interchange

A standard for interchange of  $\mu(E)$  data would address most of these concerns.

# Goals of a data interchange format

- 1 Establish a common language for transferring data between XAS experimenters, data analysis packages, web applications, journals and anything else that needs to process XAS data.
- 2 Increase the relevance and longevity of experimental data by reducing the amount of data archeology future interpretations of that data will require.
- 3 Enhance the user experience by promoting inter-operability among data acquisition systems, data analysis packages, and other applications.
- 4 Provide a mechanism for extracting and preserving a single XAS or XAS-like data set from a related experiment or from a complex data structure.

# Data and metadata

## Data

- 1 Energy axis
- 2  $\mu(E)$ , possibly normalized  $\mu(E)E$
- 3 the  $I_0$  signal
- 4 Possibly other signal channels, integration times, other signals

## Metadata

- 1 Measurement start time, end time
- 2 Information about the sample (stoichiometry, preparation method, etc)
- 3 Information about the source (source type, ring current, ring energy, etc)
- 4 Information about the optics (mono crystal. mono d-spacing, mirror types, etc)
- 5 Measurement parameters (region boundaries, integration times, etc)
- 6 ... and so on ...

# XDI: XAS Data Interchange

XDI is an *ad hoc* format loosely based on the format of e-mail and structured in a way that looks like a familiar column-data file.

The metadata is grouped into arbitrary (but useful) families with a dot (.) as a “namespace” separator. Application-specific information uses it’s own “namespace”.

```
# XDI/1.0 MX/2.0
# Beamline.name: APS 10ID
# Beamline.edge-energy: 7112.00
# Beamline.d-spacing: 3.1356
# Ring.energy: 7.00
# Source.type: undulator a
# Source.undulator-harmonic: 1
# Time.start: 2005-03-08T20:08:57
# Optics.crystal: Si 111
# Optics.harmonic-rejection: flat Rh-coated mirror
# Column.1: energy eV
# Column.2: i0
# Column.3: itrans
# Column.4: ifluor
# Column.5: irefer
# MX.Num-regions: 1
# MX.SRB: 6900
# MX.SRSS: 0.5
# MX.SPP: 0.1
# MX.Settling-time: 0
# MX.Offsets: 11408.00 11328.00 13200.00 10774.00
# MX.Gains: 8.00 7.00 7.00 9.00
###
# Fe K-edge, Lepidocrocite powder on kapton tape, RT
# 4 layers of tape
# exafs, 20 invang
#---
# energy mcs3 mcs4 mcs6 mcs5
6899.9609 48120 19430 2250 54540
6900.1421 48390 19540 2260 54860
6900.5449 48520 19610 2250 55110
6900.9678 48930 19780 2280 55650
6901.3806 48460 19590 2250 55110
(....etc....)
```

# xasCIF

xasCIF would use the syntax of CIF, the Crystallographic Information File, and a dictionary of the scientific-field-specific ontology (i.e. the list of column names and their definitions) to encode the metadata.

```
data_copper_rt01
_xas_source synchrotron
_xas_source_type "APS beamline 13ID, APS Undulator A"
_xas_source_harmonic_rejection "rhodium-coated mirror"
_xas_source_power 7.0
_xas_radiation_collimation None
_xas_radiation_monochromator "Si 111"
_xas_radiation_focusing yes
_xas_exptl_start_time 2001-06-26T22:27:31
_xas_edge_element Cu
_xas_edge_line K
_xas_edge_energy 8980.0
_pd_spec_description "Cu metal foil" #<- from the Powder CIF dictionary
_xas_exptl_ambient_temperature 293
_gse_sample_crystal_structure Cu
_xas_detector_I0 "10cm / N2"
_xas_detector_I1 "10cm / N2"
loop_
_xas_processed_energy
_xas_processed_mu
_xas_processed_monitor
8779.0000 -1.3070486 149013.70
8789.0000 -1.3006104 144864.70
8799.0000 -1.3033816 132978.70
8809.0000 -1.3059724 125444.70
8819.0000 -1.3107085 121324.70
8829.0000 -1.3138152 119447.70
8839.0000 -1.3072055 119100.70
...
```

# Pros and cons

## XDI

### Pros

- familiar format
- easy to import into many common XAS and plotting programs

### Cons

- made up by a few guys with no particular authority
- no separation of syntax and ontology

## xasCIF

### Pros

- separates syntax from ontology
- easily extensible, can handle more than one spectrum per file
- uses the highly successful CIF format

### Cons

- structurally unfamiliar to XAS folk
- cannot be imported as-is by many common XAS or plotting programs

# Getting involved

While it is clear that a well-defined, broadly-agreed-upon data interchange format is needed, it is not yet clear what that format should be.

## This is an ongoing discussion

The XASFORMAT group is open to any and all interested parties.

Please join us!

<http://millenia.cars.aps.anl.gov/mailman/listinfo/xasformat>